SIAM J. APPL. MATH.
Vol. 61, No. 4, pp. 1338–1361

VARIATIONAL RESTORATION OF NONFLAT IMAGE FEATURES: MODELS AND ALGORITHMS*

TONY CHAN† AND JIANHONG SHEN‡

Abstract. We develop both mathematical models and computational algorithms for variational denoising and restoration of nonflat image features. Nonflat image features are those that live on Riemannian manifolds, instead of on the Euclidean spaces. Familiar examples include the *orientation* feature (from optical flows or gradient flows) that lives on the unit circle $\mathbf{S}^1$, the *alignment* feature (from fingerprint waves or certain texture images) that lives on the real projective line $\mathbb{RP}^1$, and the *chromaticity* feature (from color images) that lives on the unit sphere $\mathbf{S}^2$. In this paper, we apply the variational method to denoise and restore general nonflat image features. Mathematical models for both continuous image domains and discrete domains (or *graphs*) are constructed. Riemannian objects such as *metric, distance* and *Levi–Civita connection* play important roles in the models. Computational algorithms are also developed for the resulting nonlinear equations. The mathematical framework can be applied to restoring general nonflat data outside the scope of image processing and computer vision.

Key words. variational model, total variation, denoising and restoration, nonflat features, Riemannian manifold, metric and distance, orientation, alignment, chromaticity

AMS subject classifications. Primary, 9412; Secondary, 65H10, 49M37

PII. S003613999935799X

1. Introduction.

In image analysis, restoration and denoising by variational methods are traditionally applied to gray level images or vectorial (or color) images [8, 9, 14, 22]. Such functions are "flat" or affine. By flatness or affinity, we mean that a function takes all possible values in an open set (or its closure) in a linear space.

Recently in image processing there has been increasing interest in nonflat image features that live on curved manifolds [11, 17, 18, 19, 20]. One example is the *orientation feature*, coming from optical flows or gradient flows (of gray images). For a general vector field, orientation encodes at least one-half of the whole information (and the other half is the magnitude or length), and in many applications, it is often independently a crucial image feature. The second example is the *alignment feature* which is commonly found in texture images. For example, for fingerprints, the aligning pattern of fingerprint "waves" is crucial for identification (see Perona [11]). The *chromaticity feature* of color images is another important example of nonflat features (see Tang, Sapiro, and Caselles [17, 18]). In an RGB color image, the pixel value $\mathbf{I}$ is a vector in $\mathbb{R}^3$. The magnitude $\|\mathbf{I}\|$ measures the brightness, while its trace on the unit sphere $\mathfrak{f} = \mathbf{I}/\|\mathbf{I}\|$ encodes the color saturation information, and is called the *chromaticity feature* in Tang, Sapiro, and Caselles [18]. Apparently, chromaticity lives on the unit sphere $\mathbf{S}^2$.

Mathematically, nonflat features are modeled by differential manifolds and perceptual sensitivity is measured by Riemannian metrics. Therefore, our models shall

*Received by the editors June 21, 1999; accepted for publication (in revised form) May 1, 2000; published electronically December 7, 2000. This research was supported by grants from NSF under grant DMS-9626755 and from ONR under N00014-96-1-0277.
http://www.siam.org/journals/siap/61-4/35799.html
†Department of Mathematics, UCLA, Los Angeles, CA 90095-1555 (chan@math.ucla.edu).
‡School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (jhshen@math.umn.edu).

be inevitably based on Riemannian manifolds. The complete picture has been summarized in Figure 1.

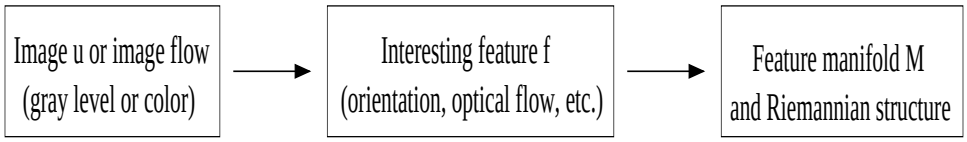

FIG. 1. *The flow chart of model setup.*

In the first step, meaningful features are often defined or extracted according to our visual system or the type of tasks we attempt to carry out. The second step demands mathematical analysis. A perceptually meaningful Riemannian metric depends on our visual system. For example, orientations in two-dimensional (2-D) images live on the unit circle. It is reasonable to assume that generally our visual system does not prefer any specific angle ($\pi/4$, say). Thus we are able to use the standard Riemannian structure (rotationally invariant) for orientations. In practice, the converse is often true: familiar manifolds usually have their natural (i.e., invariant to certain transform groups) Riemannian metrics, on which models can be conveniently based. The effectiveness can be checked and modified later by the numerical outcomes of such models.

Our starting point is the last box. That is, a perceptually meaningful Riemannian structure is assumed to be already defined.

In most applications, nonflat features obtained initially are often noisy due to some common factors, such as

(1) the noise contaminating the original images, or left there even after a denoising preprocessing;
(2) regional destruction of raw images (due to the aging of a film, etc.), or insufficiency of resolution;
(3) inefficiency of the algorithms that extract nonflat features.

As for flat features, the variational method is one of the most powerful tools to restore and denoise nonflat features. Properly designed variational models can (i) attenuate noise, (ii) preserve intrinsic singularities, and (iii) enhance the "edges" of nonflat features. For general features (flat or nonflat), an "edge" refers to a 1-D segment along which discontinuity occurs. For the orientation feature from optical flows, for example, the edges may be identical to the physical boundaries of moving bodies.

Perona has taken the pioneering step in [11] by studying a discrete variational model for orientation diffusion with applications to fingerprint analysis. Our present work extends it both theoretically and computationally, and both in depth and generality. (Almost at the same time, Sapiro and his colleagues were also developing the diffusion theory for directions [17, 18].)

Two key mathematical challenges exist for restoring nonflat features. First, it is not so obvious as in Euclidean spaces to come up with feasible energy functionals for nonflat features. Perona [11] offered some bright physical intuition for the orientation feature, but a general mathematical framework is still missing, which may delay the study of many other potentially important nonflat features. Secondly, the associated Euler–Lagrange equations or steepest descent equations are generally nonlinear due to the nonflatness of the feature manifolds and the intrinsic nonlinearity of their classical counterparts (the total variational (TV) model, for instance). Algorithms conquering

such high nonlinearity are thus crucial for successful applications of the restoration models.

The present paper is intended to study these two aspects. We establish a systematic way for constructing the energy functions for general nonflat data or features. We interpret Perona's original model in this general framework and discuss its generalizations. The mathematical framework also extends the classical TV model. Numerical evidence shows that the TV model is more successful than the L^2 model for restoring nonflat features with edges or other singularities. We also design efficient numerical algorithms to solve the nonlinear restoration equations. Numerical results highlight the success of both our mathematical models and computational algorithms.

The organization of the paper is as follows. Section 2 constructs the mathematical foundation for the work of both Perona and ours. We study the restoration of nonflat features in a general and abstract way. The focus is on the construction of the energy functionals on a Riemannian feature manifold and their resulting Euler–Lagrange equations (continuous or discrete). In section 3, we discuss specifically the restoration models for orientations and alignments, two types of nonflat image features frequently encountered in gray level image analysis. The latter half of the section is devoted to developing numerical algorithms that can efficiently solve the nonlinear restoration equations for orientations and alignments. Section 4 studies the computational issues for the chromaticity restoration. A simple two-step filtering algorithm on the sphere is established for the discrete models. Numerical examples are gathered in section 5, where we discuss applications in fingerprint images, optical flows, and color images.

2. Variational models and restoration equations. Since nonflat features do not live in flat spaces, the energy functionals and associated restoration equations depend on the "geometry" of the nonflatness. The geometry is described by Riemannian manifolds and Riemannian metrics, the basic objects that we shall work with in this section. Our goal is to construct general models for the restoration of nonflat image features.

In the two subsections, we study models for both continuous image domains and digital image domains (or graphs, more generally). The latter has the advantages of easier implementation in computers and very friendly accessibility to general readers and users. Such self-contained digitization methodology has been explored in more details recently in [3].

2.1. Models for continuous image domains. Let Ω be an image domain in $\mathbb{R}^2$. For the purpose of image analysis, Ω can be a square, a disk, or other regular domains.

Suppose that the target nonflat feature $\mathfrak{f} : \Omega \to M$ lives on an m-dimensional Riemannian manifold M. We shall call M the *feature manifold*. For example, the unit circle $\mathbf{S}^1$ is the feature manifold for orientations (or arrows with unit length), while the real projective line $\mathbb{RP}^1$ is the feature manifold for alignments (or line segments with unit length).

As pointed out in the introduction section, in practice, a nonflat feature distribution of a given image u is often corrupted by noise. Our major goal is to construct denoising and restoration models using the classical tool of variational methods.

Typically, a feature distribution $\mathfrak{f} : \Omega \to M$ is a map without overall smoothness. However, when constructing a restoration model, as is well practiced in the case of affine features, $\mathfrak{f}(\boldsymbol{p})$ is always assumed (or rather, *pretended*) to have a certain degree of regularities. For instance, we can assume that $\mathfrak{f}(\boldsymbol{p})$ belongs to certain Sobolev classes. In numerical computation, it is fortunately true that such models always

work effectively for general images (and image features). One simple reason for this is that, theoretically, we can always diffuse (or convolve with Gaussians) an image in tiny scales to mollify it, yet without introducing harmful visual effects. The next section discusses the digital model, in which the regularity of data causes no problem in both theory and computation since only discrete data are studied.

Under this setup, $\partial_x \mathfrak{f}(\boldsymbol{p})$ and $\partial_y \mathfrak{f}(\boldsymbol{p})$ are two tangent vectors in the tangent space $T_{\mathfrak{f}(\boldsymbol{p})}M$. Let $\langle\, \cdot,\, \cdot\, \rangle_{\mathfrak{f}(\boldsymbol{p})}$ denote the Riemannian inner product at $\mathfrak{f}(\boldsymbol{p})$, and let $\|\cdot\|_{\mathfrak{f}(\boldsymbol{p})}$ denote the induced norm in $T_{\mathfrak{f}(\boldsymbol{p})}M$. We shall omit the location subscript below if it is clear or insignificant in the context.

We now define the *strength* function

$$e(\mathfrak{f};\boldsymbol{p}) = \sqrt{\|\partial_x \mathfrak{f}(\boldsymbol{p})\|^2 + \|\partial_y \mathfrak{f}(\boldsymbol{p})\|^2}. \tag{2.1}$$

To gain a better understanding, we introduce the correlation matrix

$$C_{\mathfrak{f}}(\boldsymbol{p}) = \begin{bmatrix} \langle\, \partial_x \mathfrak{f},\, \partial_x \mathfrak{f}\, \rangle & \langle\, \partial_x \mathfrak{f},\, \partial_y \mathfrak{f}\, \rangle \\ \langle\, \partial_y \mathfrak{f},\, \partial_x \mathfrak{f}\, \rangle & \langle\, \partial_y \mathfrak{f},\, \partial_y \mathfrak{f}\, \rangle \end{bmatrix}.$$

Then

$$e(\mathfrak{f};\boldsymbol{p}) = \sqrt{\operatorname{trace}(C_{\mathfrak{f}}(\boldsymbol{p}))} = \sqrt{\sigma_1^2 + \sigma_2^2},$$

where σ_1 and σ_2 are the singular values of the m (the dimension of M) by 2 "matrix" $[\partial_x \mathfrak{f}, \partial_y \mathfrak{f}]$ (which is well defined under any local orthonormal frame of M). A similar formulation has also been proposed by Sapiro [15] and Sapiro and Ringach [16] when $\mathfrak{f}$ is a vectorial distribution.

The strength function defined in this way is rotationally invariant. That is, for any orthogonal transform in the image domain $\mathbf{R} : \mathbb{R}^2 \to \mathbb{R}^2$, we have

$$e(\mathfrak{f}(\mathbf{R}^{-1}\cdot);\mathbf{R}\boldsymbol{p}) = e(\mathfrak{f}(\cdot);\boldsymbol{p}),$$

since the corresponding correlation matrices are orthogonally similar to each other. This property is meaningful in real life since we do not want the results to depend on the direction in which an image is displayed before us.

The L^2 *total energy* of a feature distribution $\mathfrak{f}$ is then defined to be

$$\mathcal{E}(\mathfrak{f}) = \frac{1}{2}\int_\Omega e^2(\mathfrak{f};\boldsymbol{p})\, d\boldsymbol{p}.$$

Inspired by the affine case where Rudin and Osher [13] and Rudin, Osher, and Fatemi [14] first introduced the TV model for image restoration, we define the *total variation* to be the L^1 total energy

$$\mathcal{E}^{\mathrm{TV}}(\mathfrak{f}) = \int_\Omega e(\mathfrak{f};\boldsymbol{p})\, d\boldsymbol{p}.$$

Let d denote the metric on M induced by its Riemannian structure. That is, for any two points $\mathfrak{f}$ and $\mathfrak{g} \in M$,

$$d(\mathfrak{f},\mathfrak{g}) = \inf_\Gamma \ \operatorname{length}(\Gamma),$$

allowing Γ to go over all piecewise smooth paths that link $\mathfrak{f}$ and $\mathfrak{g}$. (We shall also allow other kind of distances later. See the discrete model in the next subsection.)

Given a noisy feature distribution $\mathfrak{f}^{(0)}(\boldsymbol{p})$, the classical restoration model associated with the L^2 energy is to solve

$$\min_{\mathfrak{f}} \mathcal{E}(\mathfrak{f}, \lambda),$$

where the fitted L^2 energy

$$\begin{aligned}\mathcal{E}(\mathfrak{f};\lambda) &= \mathcal{E}(\mathfrak{f}) + \frac{\lambda}{2}\int_\Omega d^2(\mathfrak{f}^{(0)},\mathfrak{f})\, d\boldsymbol{p} \\ &= \frac{1}{2}\int_\Omega e^2(\mathfrak{f};\boldsymbol{p})\, d\boldsymbol{p} + \frac{\lambda}{2}\int_\Omega d^2(\mathfrak{f}^{(0)},\mathfrak{f})\, d\boldsymbol{p}.\end{aligned}$$

The first term regularizes the restoration solution, while the second states the fitting constraint.

For the TV restoration, the associated *fitted TV energy* is

$$\mathcal{E}^{\mathrm{TV}}(\mathfrak{f};\lambda) = \int_\Omega e(\mathfrak{f};\boldsymbol{p})\, d\boldsymbol{p} + \frac{\lambda}{2}\int_\Omega d^2(\mathfrak{f}^{(0)},\mathfrak{f})\, d\boldsymbol{p}.$$

The constant λ above is the Lagrange multiplier for the associated constrained optimization problem

$$\text{minimize} \quad \mathcal{E}(\mathfrak{f}) \ \ (\text{or} \ \ \mathcal{E}^{\mathrm{TV}}(\mathfrak{f})) \qquad \text{subject to} \qquad \frac{1}{|\Omega|}\int_\Omega d^2(\mathfrak{f}^{(0)},\mathfrak{f}) d\boldsymbol{p} = \sigma^2,$$

where $|\Omega|$ is the area of Ω and σ the standard deviation of the noise. In the flat case, the connection between the constrained and unconstrained formulations was discussed theoretically by Chambolle and Lions [2] and numerically by Blomgren and Chan [1]. In image analysis, λ is often conveniently fixed or chosen a priori, or estimated by the gradient-projection method (Rudin, Osher, and Fatemi [14]).

To solve the above variational problems, we study the associated Euler–Lagrange equations. For the fitted L^2 energy $\mathcal{E}(\mathfrak{f};\lambda)$, the Euler–Lagrange equation can be shown to be

$$-\partial_x^*(\partial_x\mathfrak{f}) - \partial_y^*(\partial_y\mathfrak{f}) + \frac{\lambda}{2}\mathrm{grad}_{\mathfrak{f}} d^2(\mathfrak{f}^{(0)},\mathfrak{f}) = 0,$$

where, at any pixel $\boldsymbol{p} = (x, y) \in \Omega$, $\mathrm{grad}_{\mathfrak{f}} d^2(\mathfrak{f}^{(0)},\mathfrak{f})$ denotes the gradient vector of the scalar function $d^2(\mathfrak{f}^{(0)}(\boldsymbol{p}),\cdot)$ on M, and ∂_x^* and ∂_y^* denote the covariant derivatives acting on vector fields defined on $\mathfrak{f}(\Omega)$. Specifically, let ∇ be the Levi–Civita connection (a tool for differentiating vector fields on general Riemannian manifolds; see Helgason [6], for example) on the feature Riemannian manifold M. Then the covariant derivatives ∂_x^* and ∂_y^* are given explicitly by

$$\partial_x^* = \nabla_{\partial_x\mathfrak{f}}, \qquad \partial_y^* = \nabla_{\partial_y\mathfrak{f}}.$$

They can legally act on any vector field X that is defined along $\mathfrak{f}(\Omega)$ due to the locality of covariant derivatives. For instance, if M is a domain in $\mathbb{R}^m$, then we have simply

$$\partial_x^* X = \frac{\partial}{\partial x} X(\mathfrak{f}(x,y)).$$

Generally, when M is isometrically embedded in the Euclidean space $\mathbb{R}^N$ $(N \geq m)$, a vector field X on M is also one in $\mathbb{R}^N$. Since the flat derivative in $\mathbb{R}^N$,

$$\frac{\partial}{\partial x} X(\mathfrak{f}(x,y)),$$

usually overshoots from the tangent space of M, one needs an extra step of projection for the covariant derivative:

$$\partial_x^* X = \Pi_{\mathfrak{f}} \frac{\partial}{\partial x} X(\mathfrak{f}(x,y)), \tag{2.2}$$

where $\Pi_{\mathfrak{f}}$ is the orthogonal projection from $T_{\mathfrak{f}}\mathbb{R}^N$ onto the tangent space $T_{\mathfrak{f}}M$.

The Euler–Lagrange equation can be solved by the infinitesimal steepest descent marching, which leads to the *forced diffusion equation*

$$\frac{\partial \mathfrak{f}}{\partial t} = \partial_x^*(\partial_x \mathfrak{f}) + \partial_y^*(\partial_y \mathfrak{f}) - \frac{\lambda}{2}\mathrm{grad}_{\mathfrak{f}} d^2(\mathfrak{f}^{(0)}, \mathfrak{f}),$$

with the Neumann adiabatic boundary condition plus the initial condition

$$\mathfrak{f}\,\big|_{t=0} = \mathfrak{f}^0,$$

where $\mathfrak{f}^0$ can be the given raw feature distribution $\mathfrak{f}^{(0)}$ for convenience. All the terms in the diffusion equation are in the tangent space $T_{\mathfrak{f}(\boldsymbol{p})}M$. As the marching time increases, $\mathfrak{f}$ evolves to the equilibrium distribution of the functional.

Similarly, for the TV model, the Euler–Lagrange equation becomes

$$-\partial_x^*\left(\frac{\partial_x \mathfrak{f}}{e(\mathfrak{f};\boldsymbol{p})}\right) - \partial_y^*\left(\frac{\partial_y \mathfrak{f}}{e(\mathfrak{f};\boldsymbol{p})}\right) + \frac{\lambda}{2}\mathrm{grad}_{\mathfrak{f}} d^2(\mathfrak{f}^{(0)}, \mathfrak{f}) = 0.$$

And steepest descent leads to the forced anisotropic diffusion equation

$$\frac{\partial \mathfrak{f}}{\partial t} = \partial_x^*\left(\frac{\partial_x \mathfrak{f}}{e(\mathfrak{f};\boldsymbol{p})}\right) + \partial_y^*\left(\frac{\partial_y \mathfrak{f}}{e(\mathfrak{f};\boldsymbol{p})}\right) - \frac{\lambda}{2}\mathrm{grad}_{\mathfrak{f}} d^2(\mathfrak{f}^{(0)}, \mathfrak{f}).$$

Example 1 (feature manifolds embedded in $\mathbb{R}^N$). When the feature manifold M can be isometrically embedded in Euclidean space $\mathbb{R}^N$ (such as the unit circle in $\mathbb{R}^2$ and the unit sphere $\mathbf{S}^m$ in $\mathbb{R}^{m+1}$), the Levi–Civita connection is simply the Euclidean derivatives followed by a projection, as pointed out above. In this case, a feature point $\mathfrak{f}$ can be seen as an ordinary vector in $\mathbb{R}^N$, and the definition of the strength function

$$e(\mathfrak{f};\boldsymbol{p}) = \sqrt{\|\mathfrak{f}_x(\boldsymbol{p})\|^2 + \|\mathfrak{f}_y(\boldsymbol{p})\|^2}$$

now involves only the ordinary differentiations of vectors and is identical with

$$\|\nabla \mathfrak{f}\| = \sqrt{\|\nabla \mathfrak{f}_1\|^2 + \|\nabla \mathfrak{f}_2\|^2 + \cdots + \|\nabla \mathfrak{f}_N\|^2}$$

if $\mathfrak{f} = (\mathfrak{f}_1, \mathfrak{f}_2, \ldots, \mathfrak{f}_N) \in \mathbb{R}^N$. Moreover, the diminishing flow fields F of the energy functionals can be computed explicitly. When $M = \mathbf{S}^m$ is embedded in $\mathbb{R}^{m+1}$, this work has been carried out recently in Tang, Sapiro, and Caselles [17] and was also worked out independently by Osher [10]. For a general embedded feature manifold, for example, the diminishing flow field

$$F = \partial_x^*(\partial_x \mathfrak{f}) + \partial_y^*(\partial_y \mathfrak{f})$$

for the L^2 energy $\mathcal{E}$ is simply given by

$$F = \Pi_{\mathfrak{f}}\, \Delta \mathfrak{f}, \tag{2.3}$$

where Δ is the flat Laplacian on Ω and $\Pi_{\mathfrak{f}}$ is the orthogonal projection from $T_{\mathfrak{f}}\mathbb{R}^N$ onto $T_{\mathfrak{f}}M$. Spheres $\mathbf{S}^m$ in $\mathbb{R}^{m+1}$ are the most interesting to image processing. For the sphere, it was shown by Tang, Sapiro, and Caselles [17] and Osher [10] that the diminishing flow for the L^2 energy is

$$F(\mathfrak{f}) = \Delta\mathfrak{f} + \|\nabla\mathfrak{f}\|^2\mathfrak{f}, \tag{2.4}$$

and for the TV energy,

$$F^{\mathrm{TV}}(\mathfrak{f}) = \nabla\cdot\left(\frac{\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|}\right) + \|\nabla\mathfrak{f}\|\,\mathfrak{f}. \tag{2.5}$$

Osher [10] established these formulae by the method of Lagrange multiplier analysis, and Tang, Sapiro, and Caselles [17] proved them by working out explicitly the variational problem. Here we give a simpler geometric proof based on the special form of the Levi–Civita connection (2.2).

Proof of (2.4) *and* (2.5) (see Figure 2). Take the TV diminishing flow (2.5), for example. For sphere $\mathbf{S}^m$, the projection at $\mathfrak{f}$ is

$$\Pi_{\mathfrak{f}}\mathbf{v} = \mathbf{v} - (\mathbf{v}\cdot\mathfrak{f})\mathfrak{f}.$$

Hence for any distribution $\mathfrak{f}:\Omega\to\mathbf{S}^m$,

$$\begin{aligned} F^{\mathrm{TV}}(\mathfrak{f}) &= \Pi_{\mathfrak{f}}\nabla\cdot\frac{\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|} \\ &= \nabla\cdot\frac{\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|} - \left(\nabla\cdot\frac{\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|}\cdot\mathfrak{f}\right)\mathfrak{f} \\ &= \nabla\cdot\frac{\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|} - \left(\nabla\cdot\frac{\mathfrak{f}\cdot\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|} - \|\nabla\mathfrak{f}\|\right)\mathfrak{f} \\ &= \nabla\cdot\frac{\nabla\mathfrak{f}}{\|\nabla\mathfrak{f}\|} + \|\nabla\mathfrak{f}\|\,\mathfrak{f}, \end{aligned}$$

since $\mathfrak{f}\cdot\nabla\mathfrak{f} = \nabla(\mathfrak{f}^2/2) = \nabla(1/2) = 0$. □

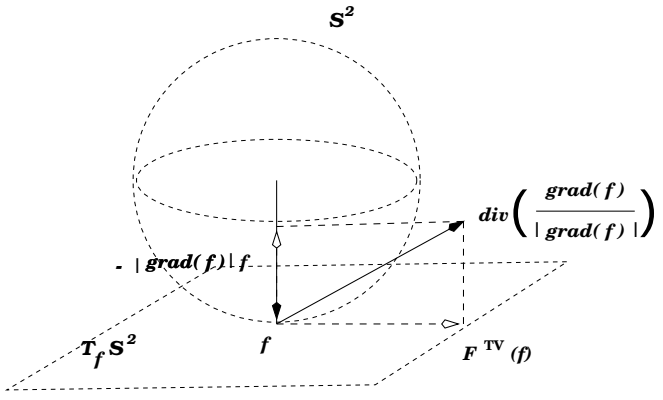

FIG. 2. *The TV diminishing flow.*

This continuous model provides a universal language for restoring nonflat features. However, the model assumes that the readers have some knowledge of modern differential geometry, concepts such as the Levi–Civita connection. It is unavoidable in this

continuous formulation since the variational formulation is global. Fortunately, there does exist a way that one can circumvent this technical difficulty. We now develop a simpler model for "digital" image domains, namely graphs, where differentiation of vector fields (i.e., Levi–Civita connection) can be avoided. Besides, the discrete models have their own computational advantages, as discussed later in this paper.

2.2. Models for discrete image domains. Since numerically it is always necessary to lay down a finite discrete grid to approximate the continuous image domain Ω, it is practically convenient and important to study directly nonflat features over a discrete grid, or more generally, a graph. This is the main task of this section. Perona took the first step in [11].

Let $G[\Omega_n]$ denote a graph supported over a finite set of nodes Ω_n ("n" stands for "numerical"), and let G denote the dictionary of edges. We write $\alpha \sim \beta$ if nodes α and β are linked by an edge. For any node $\alpha \in \Omega_n$, denote by N_α all its neighbors

$$\{\beta \in \Omega_n : \ \beta \sim \alpha\}.$$

A locally Riemannian distance (LRD) is a piecewise smooth continuous function d_l,

$$d_l : \quad M \times M \to \mathbb{R}^+ = [0, \infty),$$

such that

(1) for any $\mathfrak{f}, \mathfrak{g} \in M$, $d_l(\mathfrak{f}, \mathfrak{g}) = d_l(\mathfrak{g}, \mathfrak{f})$;
(2) $d_l(\mathfrak{f}, \mathfrak{g}) = 0$, if and only if $\mathfrak{f} = \mathfrak{g}$;
(3) $d_l(\mathfrak{f}, \mathfrak{g}) + d_l(\mathfrak{f}, \mathfrak{h}) \geq d_l(\mathfrak{g}, \mathfrak{h})$;
(4) $d_l(\mathfrak{f}, \mathfrak{g}) = d(\mathfrak{f}, \mathfrak{g}) + O(d^3(\mathfrak{f}, \mathfrak{g}))$ as $d(\mathfrak{f}, \mathfrak{g}) \to 0$.

Here d is the induced Riemannian distance defined in the previous section.

Remark 1. Conditions (1)–(3) are the basic ingredients of a distance, which authorize d_l to "legally" measure the amplitude of noise. Condition (4) restricts our attention only to those distances that are locally Riemannian. By such a constraint, the l^2 discrete model based on d_l can lead to the ordinary Laplacian in the leading term (also see the remarks at the end of this section).

Example 2 (Perona's distance). Consider the unit circle $\mathbf{S}^1$ as the concerned feature manifold. For any two points $\mathfrak{f} = e^{i\theta_1}$ and $\mathfrak{g} = e^{i\theta_2}$, define

$$d_{\text{perona}}(\mathfrak{f}, \mathfrak{g}) = \sqrt{2(1 - \cos(\theta_1 - \theta_2))}.$$

Symmetry is immediate, and the condition $d_{\text{perona}}(\mathfrak{f}, \mathfrak{g}) = 0$ requires $\cos(\theta_1 - \theta_2) = 1$, implying that $\mathfrak{f} = \mathfrak{g}$. It is also clear that the leading term of $d_{\text{perona}}(\mathfrak{f}, \mathfrak{g})$ is the Riemannian distance when $\mathfrak{f}$ and $\mathfrak{g}$ are close. Therefore, to show d_{perona} is a locally Riemannian distance, it suffices to verify the triangular relation. We leave the simple verification to our readers. This is the distance that Perona first proposed in [11] (though there it was constructed by physics).

As an example of a non-LRD, consider

$$d_{\text{line}}(\mathfrak{f}, \mathfrak{g}) = \sqrt{\frac{1 - \cos(2(\theta_1 - \theta_2))}{2}}.$$

It is not a local distance for $\mathbf{S}^1$ since $d_{\text{line}}(1, -1) = 0$. For the real projective line $\mathbb{RP}^1$, however, d_{line} is indeed an LRD (see section 3).

Let $\mathfrak{f}_\alpha$ denote the nonflat feature at a node α. For a given LRD d_l, we define the associated *strength* of a given feature distribution $\mathfrak{f}$ to be

$$e(\mathfrak{f};\alpha) = \left[\sum_{\beta\in N_\alpha} d_l^2(\mathfrak{f}_\beta,\mathfrak{f}_\alpha)\right]^{\frac{1}{2}},$$

which generalizes the continuous strength function in (2.1) or $\|\nabla\mathfrak{f}\|$ when $\mathfrak{f}$ is a flat function. Define the l^2 total energy and the *total variation* to be

$$\mathcal{E}(\mathfrak{f}) = \sum_{\alpha\in\Omega_n} \frac{1}{2}e^2(\mathfrak{f};\alpha) \qquad \text{and} \qquad \mathcal{E}^{\mathrm{TV}}(\mathfrak{f}) = \sum_{\alpha\in\Omega_n} e(\mathfrak{f};\alpha).$$

For a given feature distribution $\mathfrak{f}^{(0)}$ and some weight constant λ, define the *fitted* l^2 total energy

$$\mathcal{E}(\mathfrak{f};\lambda) = \mathcal{E}(\mathfrak{f}) + \lambda \sum_{\alpha\in\Omega_n} \frac{1}{2} d_l^2(\mathfrak{f}_\alpha^{(0)},\mathfrak{f}_\alpha),$$

and the *fitted* TV total energy

$$\mathcal{E}^{\mathrm{TV}}(\mathfrak{f};\lambda) = \mathcal{E}^{\mathrm{TV}}(\mathfrak{f}) + \lambda \sum_{\alpha\in\Omega_n} \frac{1}{2} d_l^2(\mathfrak{f}_\alpha^{(0)},\mathfrak{f}_\alpha).$$

To establish the diffusion equations, we first compute the gradient fields for the total energies. For the l^2 total energy,

$$\frac{\partial\mathcal{E}(\mathfrak{f})}{\partial\mathfrak{f}_\alpha} = \sum_{\beta\in N_\alpha} \frac{\partial}{\partial\mathfrak{f}_\alpha} d_l^2(\mathfrak{f}_\alpha,\mathfrak{f}_\beta); \tag{2.6}$$

for the TV total energy,

$$\frac{\partial\mathcal{E}^{\mathrm{TV}}(\mathfrak{f})}{\partial\mathfrak{f}_\alpha} = \sum_{\beta\in N_\alpha} \left[\frac{\partial}{\partial\mathfrak{f}_\alpha} d_l^2(\mathfrak{f}_\alpha,\mathfrak{f}_\beta)\right] \left(\frac{1/e(\mathfrak{f};\alpha) + 1/e(\mathfrak{f};\beta)}{2}\right). \tag{2.7}$$

Hence the corresponding forced diffusion equations are

$$\frac{d\mathfrak{f}_\alpha}{dt} = -\sum_{\beta\in N_\alpha} \frac{\partial}{\partial\mathfrak{f}_\alpha} d_l^2(\mathfrak{f}_\alpha,\mathfrak{f}_\beta) - \frac{\lambda}{2}\frac{\partial}{\partial\mathfrak{f}_\alpha} d_l^2(\mathfrak{f}_\alpha^{(0)},\mathfrak{f}_\alpha); \tag{2.8}$$

$$\frac{d\mathfrak{f}_\alpha}{dt} = -\sum_{\beta\in N_\alpha} \left[\frac{\partial}{\partial\mathfrak{f}_\alpha} d_l^2(\mathfrak{f}_\alpha,\mathfrak{f}_\beta)\right] \left(\frac{1/e(\mathfrak{f};\alpha) + 1/e(\mathfrak{f};\beta)}{2}\right) - \frac{\lambda}{2}\frac{\partial}{\partial\mathfrak{f}_\alpha} d_l^2(\mathfrak{f}_\alpha^{(0)},\mathfrak{f}_\alpha). \tag{2.9}$$

Notice that here our *starting point* is the locally Riemannian distance d_l. In his work on orientation diffusion, Perona [11] started with a choice of the energy, as inspired by physics. For general nonflat feature manifolds, the distance function d_l (and the special case of $d_l = d$, the Riemannian distance) seems to be more fundamental an element than the energy itself.

As an example, we now apply the discrete models to the restoration of *chromaticity*. The continuous diffusion model without the fitting constraint was first discussed in Tang, Sapiro, and Caselles [18]. Our discrete model has the following advantages:

(1) unlike the continuous model, one does not need to choose a numerical discretization scheme (often complicated) for the spatial derivatives; (2) the discrete model rigorously diminishes the energy function, while in the continuous model, the numerical PDEs only asymptotically (i.e., as the grid size tends to zero) diminish the continuous energy functionals.

Example 3 (restoration of the chromaticity feature). The 2-D sphere $\mathbf{S}^2$ models the *chromaticity* feature of color images. Let $\mathbf{I}(\boldsymbol{p})$ denote the RGB vector value at pixel $\boldsymbol{p}$, which can be perceptually separated into the *brightness component* $B = \|\mathbf{I}(\boldsymbol{p})\|$ and the *chromaticity component* $\mathfrak{f}(\boldsymbol{p}) = \mathbf{I}(\boldsymbol{p})/B$. $\mathfrak{f}(\boldsymbol{p})$ records the saturation degree of colors. We now apply to chromaticity restoration the discrete models developed above. Take the TV model (2.9), for example.

Embed $\mathbf{S}^2$ in $\mathbb{R}^3$ and take d_l to be the embedded Euclidean distance, i.e.,

$$d_l(\mathfrak{f},\mathfrak{g}) := \|\mathfrak{f}-\mathfrak{g}\|_{\mathbb{R}^3} = \sqrt{(\mathfrak{f}-\mathfrak{g})^2} \qquad \text{for any} \quad \mathfrak{f},\mathfrak{g}\in\mathbf{S}^2.$$

Apparently, d_l is an LRD. Let $\nabla_{\mathfrak{f}}$ denote the gradient for $\mathfrak{f}\in\mathbb{R}^3$ and $\partial/\partial\mathfrak{f}$ the gradient on $\mathbf{S}^2$. Then for any scalar function $G(\mathfrak{f})$ on $\mathbb{R}^3$,

$$\frac{\partial}{\partial\mathfrak{f}}G(\mathfrak{f}) = \Pi_{\mathfrak{f}}\nabla_{\mathfrak{f}}G(\mathfrak{f}).$$

Therefore, for any fixed feature point $\mathfrak{g}\in\mathbf{S}^2$,

$$\begin{aligned}\frac{\partial}{\partial\mathfrak{f}}d_l^2(\mathfrak{f},\mathfrak{g}) &= \Pi_{\mathfrak{f}}\nabla_{\mathfrak{f}}(\mathfrak{f}-\mathfrak{g})^2\\ &= 2\Pi_{\mathfrak{f}}(\mathfrak{f}-\mathfrak{g}) = -2\Pi_{\mathfrak{f}}(\mathfrak{g})\\ &= -2(\mathfrak{g}-(\mathfrak{g}\cdot\mathfrak{f})\mathfrak{f}).\end{aligned}$$

The TV model (2.9) is thus explicitly given by

$$\frac{d\mathfrak{f}_\alpha}{dt} = \sum_{\beta\in N_\alpha}\Pi_{\mathfrak{f}_\alpha}(\mathfrak{f}_\beta)\left(\frac{1}{e(\mathfrak{f};\alpha)}+\frac{1}{e(\mathfrak{f};\beta)}\right)+\lambda\Pi_{\mathfrak{f}_\alpha}(\mathfrak{f}_\alpha^{(0)}). \tag{2.10}$$

There are no spatial derivatives in the last equation! Thus, one only needs to take care of the time discretization, which is easy to handle with the *geodesic marching scheme* we propose in section 4. It is also plain to see that the fitted diminishing flow on the right-hand side is indeed on the tangent plane at $\mathfrak{f}_\alpha$, which is crucial to prevent $\mathfrak{f}_\alpha$ from wandering away from the feature manifold infinitesimally. This completes our example. We will return to it later in section 4 for its detailed numerical implementation.

More generally, suppose the feature manifold M is isometrically embedded in $\mathbb{R}^N$. Let $\Pi_{\mathfrak{f}}$ denote the orthogonal projection from $T_{\mathfrak{f}}\mathbb{R}^N$ onto the tangent plane $T_{\mathfrak{f}}M$. If we choose d_l to be the straight line distance in $\mathbb{R}^N$ (but restricted on M), then the TV restoration equation is given by

$$\frac{d\mathfrak{f}_\alpha}{dt} = \sum_{\beta\in N_\alpha}\Pi_{\mathfrak{f}_\alpha}(\mathfrak{f}_\beta-\mathfrak{f}_\alpha)\left(\frac{1}{e(\mathfrak{f};\alpha)}+\frac{1}{e(\mathfrak{f};\beta)}\right)+\lambda\Pi_{\mathfrak{f}_\alpha}(\mathfrak{f}_\alpha^{(0)}-\mathfrak{f}_\alpha). \tag{2.11}$$

Especially if M is the level set of a nonsingular (i.e., with nonzero gradient everywhere) function $\phi(\mathfrak{f})$ in R^{m+1}: $\phi=0$, then

$$\Pi_{\mathfrak{f}}\mathbf{v} = \mathbf{v}-(\mathbf{v}\cdot\phi_{\mathfrak{f}})\frac{\phi_{\mathfrak{f}}}{\|\phi_{\mathfrak{f}}\|^2},$$

where $\phi_{\mathfrak{f}}$ denotes the gradient of ϕ and $\mathbf{v} \in T_{\mathfrak{f}}\mathbb{R}^{m+1}$. In the continuous case, Osher [10] first proposed to consider the level-set feature manifolds (and indeed, both $\mathbf{S}^1$ and $\mathbf{S}^2$ are level-set manifolds).

Remark 2.

(1) In the classical case of flat features, the above digitization methodology is also discussed more systematically in [3]. The transition from continuous variational problems and Euler–Lagrange equations to the equivalent but self-contained counterparts on discrete graphs is very much inspired by the *spectral graph theory* [5], in which the graph Laplacians and graph heat equations are generated exactly in the same fashion.

(2) For graph *total variation*, the right-hand side of formula (2.7) generalizes the continuous curvature term frequently appearing in image analysis [13, 14]. Therefore, on a general graph, formula (2.9) defines a nonlinear adaptive diffusion, similar to the anisotropic diffusions in the classical literature [12]. Such an adaptive diffusion is crucial for preserving or enhancing singular features during the denoising and restoration process.

3. Restoration of orientation and alignment: Equations and algorithms. In this section, based on the general mathematical models constructed above, we study the restoration of two classes of nonflat features frequently encountered in gray level image analysis, namely, the orientation feature and the alignment feature.

3.1. Restoration equations for orientation. Define the sawtooth function on $\mathbb{R}$:

$$D(x) = \begin{cases} |x|, & |x| \le \pi, \\ 2\pi \text{ periodic extension} & \text{for general } x. \end{cases}$$

Then $D(\theta_1 - \theta_2)$ is the Riemannian distance $d = d(e^{i\theta_1}, e^{\theta_2})$, or simply, $d(\theta_1, \theta_2) = D(\theta_1 - \theta_2)$. It measures the length of the shortest arc on $\mathbf{S}^1$ that links the two feature points. Recall that in the previous example, we discussed Perona's locally Riemannian distance

$$d_{\text{perona}}(\theta_1, \theta_2) = \sqrt{2(1 - \cos(\theta_1 - \theta_2))} = 2\left|\sin\frac{\theta_1 - \theta_2}{2}\right|.$$

The fact that it is an LRD is better seen geometrically through Figure 3, which shows that d_{perona} measures the Euclidean distance of the two feature points $e^{i\theta_1}$ and $e^{i\theta_2}$ in $\mathbb{R}^2$ (instead of measuring along $\mathbf{S}^1$).

One advantage of d_{perona} over D is the smoothness of d^2_{perona}, on which we take derivatives in the restoration formulae (2.6)–(2.9). Remarkably, d_{perona} is also rotationally invariant. Therefore, in what follows, we shall discuss only this distance function.

We now derive the parametric restoration equations for d_{perona}. Parameterize the circle by $\theta \to e^{i\theta}$. Then the tangent vector $\partial/\partial\theta$ (or, equivalently, $ie^{i\theta}$) is an orthonormal basis for the tangent space at each point. All vectors appearing in the

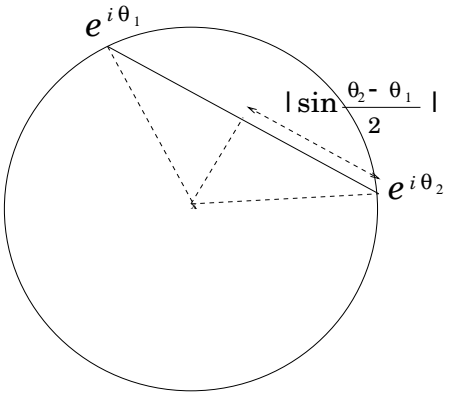

FIG. 3. *The Perona distance.*

diffusion equations are expressible using this basis. Assume $\mathfrak{f}_\alpha = e^{i\theta_\alpha}$. Then

$$\frac{\partial \mathfrak{f}_\alpha}{\partial t} = \frac{\partial \theta_\alpha}{\partial t}\frac{\partial}{\partial \theta_\alpha},$$
$$\frac{\partial d^2_{\text{perona}}(\mathfrak{f}_\alpha, \mathfrak{f}_\beta)}{\partial \mathfrak{f}_\alpha} = 2\frac{\partial(1-\cos(\theta_\alpha - \theta_\beta))}{\partial \theta_\alpha}\frac{\partial}{\partial \theta_\alpha} = 2\sin(\theta_\alpha - \theta_\beta)\frac{\partial}{\partial \theta_\alpha},$$
$$\frac{\partial d^2_{\text{perona}}(\mathfrak{f}_\alpha, \mathfrak{f}^{(0)}_\alpha)}{\partial \mathfrak{f}_\alpha} = 2\frac{\partial(1-\cos(\theta_\alpha - \theta^{(0)}_\alpha))}{\partial \theta_\alpha}\frac{\partial}{\partial \theta_\alpha} = 2\sin(\theta_\alpha - \theta^{(0)}_\alpha)\frac{\partial}{\partial \theta_\alpha}.$$

Then the fitted l^2 diffusion equation (2.8) becomes

$$\frac{d\theta_\alpha}{dt} = \sum_{\beta \in N_\alpha} 2\sin(\theta_\beta - \theta_\alpha) + \lambda \sin(\theta^{(0)}_\alpha - \theta_\alpha), \tag{3.1}$$

which is exactly the equation that Perona first studied in [11]. Similarly, the fitted TV diffusion equation is

$$\frac{d\theta_\alpha}{dt} = \sum_{\beta \in N_\alpha} \sin(\theta_\beta - \theta_\alpha)\left(\frac{1}{e(\theta;\alpha)} + \frac{1}{e(\theta;\beta)}\right) + \lambda \sin(\theta^{(0)}_\alpha - \theta_\alpha), \tag{3.2}$$

where, according to the previous section, the strength is

$$e(\theta;\alpha) = \left[\sum_{\beta \in N_\alpha} 2(1-\cos(\theta_\beta - \theta_\alpha))\right]^{\frac{1}{2}}.$$

Remark 3. Notice that the parametric (3.1) and (3.2) are equivalent to (2.8) and (2.9). This is one of the main advantages of the discrete (graph) model—being less sensitive to the discontinuities in the raw data $\theta^{(0)}$. In fact, a graph is "blind" to discontinuities since there is no infinitesimal distance. For the continuous model, the parametric equations (i.e., on θ) are not equivalent to the corresponding embedded equations (i.e., on $\mathfrak{f}$, treated as a vector in $\mathbb{R}^2$), when the given raw data $\theta^{(0)}$ has discontinuities.

Equations (3.1) and (3.2) generalize the classical equations on flat features. To see it, consider a smooth orientation distribution θ_α, that is,

$$\theta_\beta - \theta_\alpha = O(\epsilon)$$

for all $\beta \in N_\alpha$ and some small characteristic parameter ϵ. Then

$$\sum_{\beta \in N_\alpha} \sin(\theta_\beta - \theta_\alpha) = \sum_{\beta \in N_\alpha} (\theta_\beta - \theta_\alpha) + O(\epsilon^3).$$

The leading term is exactly the flat Laplacian on θ. This is generally true for any choice of LRDs because of the last condition in its definition list.

Similarly, the strength function can be shown to be

$$e(\theta; \alpha) = \left[\sum_{\beta \in N_\alpha} (\theta_\beta - \theta_\alpha)^2 \right]^{\frac{1}{2}} (1 + O(\epsilon^2)).$$

For example, for a typical internal pixel $\alpha = (i, j)$ in an image, its neighbors are

$$\beta_u = (i, j-1), \quad \beta_d = (i, j+1), \quad \beta_l = (i-1, j), \quad \beta_r = (i+1, j).$$

Thus the leading term of the strength function is simply the finite difference approximation to the magnitude of the gradient

$$\sqrt{2}h\|\nabla\theta\| = \sqrt{2}h \left[\left(\frac{\partial\theta}{\partial x} \right)^2 + \left(\frac{\partial\theta}{\partial y} \right)^2 \right]^{\frac{1}{2}},$$

where h is the discretization step size.

3.2. Restoration equations for alignment. For alignment, the direction of a line segment is unimportant (see Figure 4). This amounts to saying that the feature manifold is the real projective line $\mathbb{RP}^1$.

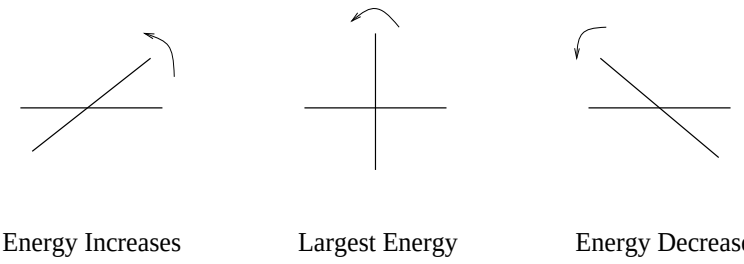

FIG. 4. *The alignment feature.*

A line segment in the direction of θ is denoted by l_θ. Thus l_θ and $l_{\theta+\pi}$ denote the same line segment. The Riemannian metric on $\mathbb{RP}^1$ is given by

$$d(l_{\theta_1}, l_{\theta_2}) = |\theta_1 - \theta_2|$$

whenever $|\theta_1 - \theta_2| < \pi/2$.

The line distance mentioned in section 2 is given by

$$d_{\text{line}}(l_{\theta_1}, l_{\theta_2}) = \sqrt{\frac{1 - \cos(2(\theta_1 - \theta_2))}{2}} = |\sin(\theta_1 - \theta_2)|.$$

d_{line} is the pull-back of d_{perona} under the circle representation of $\mathbb{RP}^1$: $l_\theta \to e^{i2\theta}$. In fact, any LRD d_l on the circle can be pulled back to $\mathbb{RP}^1$ by defining

$$d_l^{\mathbb{RP}^1}(l_{\theta_1}, l_{\theta_2}) = \frac{1}{2} d_l(e^{i2\theta_1}, e^{i2\theta_2}).$$

In what follows, we shall work only with d_{line} for the same reason of d_{perona}.

Similar to the orientation diffusion, we obtain the fitted l^2 diffusion equations for d_{line}

$$\frac{d\theta_\alpha}{dt} = \sum_{\beta\in N_\alpha} \sin(2\theta_\beta - 2\theta_\alpha) + \frac{\lambda}{2}\sin(2\theta_\alpha^{(0)} - 2\theta_\alpha). \tag{3.3}$$

The fitted TV diffusion equation is

$$\frac{d\theta_\alpha}{dt} = \sum_{\beta\in N_\alpha} \frac{1}{2}\sin(2\theta_\beta - 2\theta_\alpha)\left(\frac{1}{e(\theta;\alpha)} + \frac{1}{e(\theta;\beta)}\right) + \frac{\lambda}{2}\sin(2\theta_\alpha^{(0)} - 2\theta_\alpha), \tag{3.4}$$

where the local strength is given by

$$e(\theta;\alpha) = \left[\sum_{\beta\in N_\alpha} \frac{1-\cos(2\theta_\beta - 2\theta_\alpha)}{2}\right]^{\frac{1}{2}}.$$

We shall discuss only these two equations in the following. Our later numerical experiment shows that the alignment restoration model is especially useful for fingerprint processing (also see Perona [11]).

3.3. Algorithms. The parametric restoration equations (3.1)–(3.4) are all nonlinear. The good news from the discrete model is that the steady equations are all algebraic. Various existing solvers for nonlinear algebraic equations can thus make their contributions, though careful modifications should be made.

In this section, we discuss two numerical algorithms. We first study the l^2 restoration equations in detail, and then we point out how to make suitable adjustment for the TV restoration equations. Let Ω_n denote the graph image domain. For the purpose of image processing, one can be happy with the standard rectangular grid, where each (internal) node $\alpha = (i,j)$ has four neighboring nodes $\beta = (i, j\pm 1), (i\pm 1, j)$ such that $\beta \sim \alpha$. Considering that our models and algorithms may be applicable even out of the scope of image processing, in what follows we shall state the algorithms using the abstract notation for nodes: $\alpha, \beta, \ldots$.

We now try to solve directly the equilibrium restoration equations

$$2\sum_{\beta\sim\alpha} \sin(\theta_\beta - \theta_\alpha) + \lambda\sin(\phi_\alpha - \theta_\alpha) = 0, \qquad \alpha \in \Omega_n. \tag{3.5}$$

Here ϕ is the given raw (noisy) angle distribution. The associated total energy is given by

$$\mathcal{E}(\theta;\lambda) = 2\sum_{\beta\sim\gamma}(1-\cos(\theta_\beta - \theta_\gamma)) + \lambda\sum_{\alpha\in\Omega_n}(1-\cos(\phi_\alpha - \theta_\alpha)),$$

where $\beta\sim\gamma$ denotes all distinct edges.

3.3.1. The nonlinear Gauss–Seidel method. We first propose a nonlinear Gauss–Seidel method. Choose an initial guess θ^0 ($\theta^0 = \phi$, say) and assign an arbitrary ordering for all nodes:

$$\cdots < \beta < \alpha < \gamma < \cdots.$$

For convenience, for each node α define

$$N_\alpha^- := \{\beta \in N_\alpha : \beta < \alpha\}, \qquad N_\alpha^+ := \{\beta \in N_\alpha : \beta > \alpha\}.$$

Let k denote the global updating clock and α a certain node. Suppose that in the beginning of step $k|\alpha$, we have the following data available:

$$\theta_\beta^k : \beta < \alpha; \quad \theta_\alpha^{k-1}; \quad \theta_\beta^{k-1} : \beta > \alpha.$$

We try to update θ_α^{k-1} to θ_α^k by the end of step $k|\alpha$.

Set $x = \theta_\alpha^k$. According to the restoration equation (3.5), we demand that

$$2 \sum_{\beta \in N_\alpha^+} \sin(\theta_\beta^{k-1} - x) + 2 \sum_{\beta \in N_\alpha^-} \sin(\theta_\beta^k - x) + \lambda \sin(\phi_\alpha - \theta_\alpha^{k-1}) = 0. \tag{3.6}$$

Define

$$C = 2 \sum_{\beta \in N_\alpha^+} \cos(\theta_\beta^{k-1}) + 2 \sum_{\beta \in N_\alpha^-} \cos(\theta_\beta^k), \tag{3.7}$$

$$S = 2 \sum_{\beta \in N_\alpha^+} \sin(\theta_\beta^{k-1}) + 2 \sum_{\beta \in N_\alpha^-} \sin(\theta_\beta^k), \tag{3.8}$$

$$r = \sqrt{C^2 + S^2}, \quad \omega = \angle(C, S). \tag{3.9}$$

Here $\angle(a, b)$ denotes the angle in the direction of (a, b). Then (3.6) simplifies to

$$\sin(\omega - x) = -\frac{\lambda}{r} \sin(\phi_\alpha - \theta_\alpha^{k-1}),$$

which gives (mod 2π)

$$x = \omega + \arcsin\left[\frac{\lambda}{r} \sin(\phi_\alpha - \theta_\alpha^{k-1})\right], \quad \text{or} \quad x = \omega - \arcsin\left[\frac{\lambda}{r} \sin(\phi_\alpha - \theta_\alpha^{k-1})\right] - \pi.$$

Here we take the principle range of arcsin to be $[-\pi/2,\ \pi/2]$. To determine which one we should choose, consider the updated orientation distribution

$$\theta' = (\theta_\beta^k : \beta < \alpha, \quad x, \quad \theta_\beta^{k-1} : \beta > \alpha).$$

We select the x that gives the least total l^2 energy

$$\mathcal{E}(\theta') = 2 \sum_{\beta \sim \gamma} (1 - \cos(\theta'_\beta - \theta'_\gamma)).$$

A simple computation eventually shows that we should take

$$\theta_\alpha^k = x = \omega + \arcsin\left[\frac{\lambda}{r} \sin(\phi_\alpha - \theta_\alpha^{k-1})\right]. \tag{3.10}$$

This completes the $k|\alpha$-step of the nonlinear Gauss–Seidel updating.

Gauss–Seidel is simple, local, and reliable, but with a relatively slow convergence. Faster algorithms can be constructed using the linearization technique.

3.3.2. Linearized iteration method. Let $\operatorname{sinc}(x)$ denote the function $\sin x/x$. Then (3.5) can be written as

$$\sum_{\beta\sim\alpha}(\theta_\beta-\theta_\alpha)2\operatorname{sinc}(\theta_\beta-\theta_\alpha)+\lambda(\phi_\alpha-\theta_\alpha)\operatorname{sinc}(\phi_\alpha-\theta_\alpha)=0. \tag{3.11}$$

The *linearized iteration method* (LIM) generates the iteration as follows. Take a guess θ^0 (taken to be ϕ, say) for the initial step 0. We update θ^{k-1} to θ^k by solving a linear algebraic system. Define

$$a_{\beta\gamma}=2\operatorname{sinc}(\theta_\beta^{k-1}-\theta_\gamma^{k-1}) \quad \text{if} \quad \beta\sim\gamma; \tag{3.12}$$

$$c_\alpha=\lambda\operatorname{sinc}(\phi_\alpha-\theta_\alpha^{k-1}) \quad \text{for all} \quad \alpha\in\Omega_n. \tag{3.13}$$

Based on (3.11), for each α we require that

$$\sum_{\beta\sim\alpha}a_{\alpha\beta}(\theta_\beta^k-\theta_\alpha^k)+c_\alpha(\phi_\alpha-\theta_\alpha^k)=0.$$

This leads to an ordinary linear system

$$A\theta^k=\mathbf{b}$$

with $\mathbf{b}=(c_\alpha\phi_\alpha:\alpha\in\Omega_n)^{\mathrm{T}}$ and $A=(A_{\alpha\beta}:\alpha,\beta\in\Omega_n)$:

$$A_{\alpha\alpha}=c_\alpha+\sum_{\beta\sim\alpha}a_{\alpha\beta},$$

$$A_{\alpha\beta}=\begin{cases}-a_{\alpha\beta} & \text{if} \quad \beta\in N_\alpha,\\ 0 & \text{else.}\end{cases}$$

By such linearization, the nonlinear updating can be achieved using any fast numerical linear algebra solver.

3.3.3. Algorithms for the TV restoration equations. With the algorithms ready for the fitted l^2 energy, it is easy to modify each one separately for the fitted TV energy. We summarize them briefly below.

The nonlinear Gauss–Seidel method. The modification is done by computing the local strength e's explicitly without involving x. Therefore, (3.9) holds with a new pair of constants (C,S). The discussion there is also valid. For the flat case, the idea also appeared in Chan and Vese [4].

The LIM. The sin-sinc linearization technique extends to the TV model by modifying the constant $a_{\alpha\beta}$ to

$$a_{\alpha\beta}=\operatorname{sinc}(\theta_\beta^{k-1}-\theta_\alpha^{k-1})\left[\frac{1}{e(\theta^{k-1};\alpha)}+\frac{1}{e(\theta^{k-1};\beta)}\right].$$

This generalizes the algorithm for the classical flat case

$$0=-\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right)+\lambda(u-u^0),$$

where Vogel and Oman [21] proposed the linear iteration scheme

$$0 = -\nabla \cdot \left(\frac{\nabla u^k}{|\nabla u^{k-1}|} \right) + \lambda (u^k - u^0).$$

Remark 4. In practice, the local strength function e is usually modified to

$$e_a = \sqrt{a^2 + e^2},$$

for some small parameter a ($a = 10^{-4}$, for example). This modification avoids the zero denominator in all the above TV formulae; on the other hand, it produces an intermediate effect between the TV energy and the l^2 energy. If a is large enough, then this modified TV model gets closer to the l^2 energy. Such a regularization technique is commonly practiced in the classical literature (see [7], for example).

4. Restoration of chromaticity: Algorithms and discussions. In this section, we discuss the restoration models for the chromaticity feature living on the unit sphere $\mathbf{S}^2$ and the related computational issues. We shall discuss the computational issues only for the discrete models developed in section 2. For the work on the continuous pure diffusion model (i.e., without the fitting term), see Tang, Sapiro, and Caselles [17, 18].

4.1. The continuous models. Let $\mathfrak{f} : \Omega \to \mathbf{S}^2$ be a given raw noisy chromaticity distribution. In section 2, we have discussed the continuous restoration model corresponding to the L^2 and TV regularization energies:

$$(L^2) \qquad \mathfrak{f}_t = \Delta \mathfrak{f} + \|\nabla \mathfrak{f}\|^2 \mathfrak{f} + \lambda \Pi_{\mathfrak{f}} (\mathfrak{f}^{(0)} - \mathfrak{f}),$$

$$(\mathrm{TV}) \qquad \mathfrak{f}_t = \nabla \cdot \left(\frac{\nabla \mathfrak{f}}{\|\nabla \mathfrak{f}\|} \right) + \|\nabla \mathfrak{f}\| \mathfrak{f} + \lambda \Pi_{\mathfrak{f}} (\mathfrak{f}^{(0)} - \mathfrak{f}).$$

Here we have chosen the distance function for the fitting term to be the restricted Euclidean distance in $\mathbb{R}^3$ (not the geodesic distance on $\mathbf{S}^2$). The evolutions can start with the initial condition

$$\mathfrak{f}\big|_{t=0} = \mathfrak{f}^{(0)},$$

for instance.

There are two practical problems with these continuous models. Take the TV model, for instance. First, if a numerical scheme is sought, we must take special care for the diminishing flow term

$$F^{\mathrm{TV}}(\mathfrak{f}) = \nabla \cdot \left(\frac{\nabla \mathfrak{f}}{\|\nabla \mathfrak{f}\|} \right) + \|\nabla \mathfrak{f}\| \, \mathfrak{f}$$

to make sure that the numerical $F^{\mathrm{TV}}(\mathfrak{f})$ does lie on the tangent plane $T_{\mathfrak{f}} S^2$. If the distribution $\mathfrak{f}$ is smooth, this tangency condition causes no serious problem since it is in any case satisfied to the leading order. But our task is to denoise and restore feature distributions with poor regularities! Discontinuities and singularities must be allowed. In such a realistic situation, typically, the numerical diminishing flow F^{TV} overshoots the tangent plane $T_{\mathfrak{f}} \mathbf{S}^2$. This can, of course, be fixed by artificially projecting it back onto the tangent plane. But by doing so, we introduce a new degree of uncertainty regarding the precise diminishing effect of the algorithm.

Besides this overshooting problem, another issue is that we lack a rigorous explanation for why the numerical evolution truly minimizes the continuous energy functional $\mathcal{E}^{\mathrm{TV}}(\mathfrak{f};\lambda)$ on $\mathbf{S}^2$. The gap between continuous energy functionals and numerical PDEs cannot be smoothed away easily because of the discretization error.

These are the intrinsic problems with the continuous model. They disappear in the discrete formulation, since the latter starts with discrete energies and results in algebraic equations (on $\mathbf{S}^2$), which need no numerical discretization for spatial derivatives.

4.2. The discrete models. Let $\mathfrak{f}^{(0)}:\Omega_n\to\mathbf{S}^2$ be the given raw discrete chromaticity distribution. By taking d_l to be the restricted Euclidean distance, we have established in section 2 the discrete restoration equations corresponding to the l^2 and TV energies:

$$(l^2)\qquad \frac{d\mathfrak{f}_\alpha}{dt}=2\sum_{\beta\in N_\alpha}\Pi_{\mathfrak{f}_\alpha}\mathfrak{f}_\beta+\lambda\Pi_{\mathfrak{f}_\alpha}\mathfrak{f}_\alpha^{(0)},\qquad \alpha\in\Omega_n,$$

$$(\mathrm{TV})\quad \frac{d\mathfrak{f}_\alpha}{dt}=2\sum_{\beta\in N_\alpha}\Pi_{\mathfrak{f}_\alpha}\mathfrak{f}_\beta\left(\frac{1}{e(\mathfrak{f};\alpha)}+\frac{1}{e(\mathfrak{f};\beta)}\right)+\lambda\Pi_{\mathfrak{f}_\alpha}\mathfrak{f}_\alpha^{(0)},\qquad \alpha\in\Omega_n.$$

Define the weight

$$w_{\alpha\beta}(\mathfrak{f})=\frac{1}{e(\mathfrak{f};\alpha)}+\frac{1}{e(\mathfrak{f};\beta)}$$

for any $\beta\sim\alpha$. Then $w_{\alpha\beta}=w_{\beta\alpha}$. Rewrite the above equations using the linearity of the projection operator

$$\frac{d\mathfrak{f}_\alpha}{dt}=\Pi_{\mathfrak{f}_\alpha}\left(\sum_{\beta\in N_\alpha}2\mathfrak{f}_\beta+\lambda\mathfrak{f}_\alpha^{(0)}\right)=:F_\alpha(\mathfrak{f}),\qquad \alpha\in\Omega_n,$$

$$\frac{d\mathfrak{f}_\alpha}{dt}=\Pi_{\mathfrak{f}_\alpha}\left(\sum_{\beta\in N_\alpha}w_{\alpha\beta}\mathfrak{f}_\beta+\lambda\mathfrak{f}_\alpha^{(0)}\right)=:F_\alpha^{\mathrm{TV}}(\mathfrak{f}),\qquad \alpha\in\Omega_n.$$

The equilibrium restoration equations are thus given by

$$F_\alpha(\mathfrak{f})=\Pi_{\mathfrak{f}_\alpha}\left(\sum_{\beta\in N_\alpha}2\mathfrak{f}_\beta+\lambda\mathfrak{f}_\alpha^{(0)}\right)=0,\qquad \alpha\in\Omega_n,\tag{4.1}$$

$$F_\alpha^{\mathrm{TV}}(\mathfrak{f})=\Pi_{\mathfrak{f}_\alpha}\left(\sum_{\beta\in N_\alpha}w_{\alpha\beta}\mathfrak{f}_\beta+\lambda\mathfrak{f}_\alpha^{(0)}\right)=0,\qquad \alpha\in\Omega_n.\tag{4.2}$$

Here 0 denotes the zero vector in $T_{\mathfrak{f}}\mathbf{S}^2$. These are "algebraic" equations on the chromaticity sphere $\mathbf{S}^2$.

We now discuss the numerical implementations of both the evolution equations and the steady algebraic equations.

4.2.1. Geodesic marching for the discrete evolution equation. Take the TV model

$$\frac{df_\alpha}{dt}=F_\alpha^{\mathrm{TV}}(\mathfrak{f}),\qquad \alpha\in\Omega_n,$$

for example. Set $\mathfrak{f}_\alpha^n = \mathfrak{f}_\alpha(n\Delta t)$. The classical Euler method is a flat marching

$$\tilde{\mathfrak{f}}_\alpha^n = \mathfrak{f}_\alpha^{n-1} + \Delta t F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1}),$$

followed by a projection (see Figure 5)

$$\mathfrak{f}_\alpha^n = \frac{\tilde{\mathfrak{f}}_\alpha^n}{\|\tilde{\mathfrak{f}}_\alpha^n\|}.$$

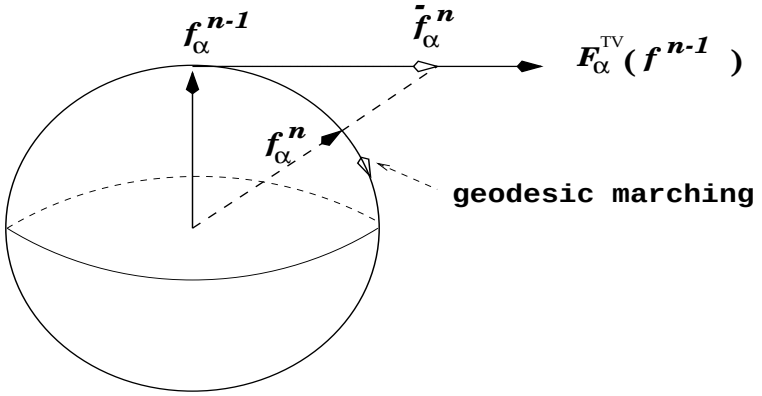

FIG. 5. *A flat marching followed by a projection, and the geodesic marching.*

To compress these two steps into a single one, we propose the *geodesic marching*:

$$\mathfrak{f}_\alpha^n = \cos\left[\Delta t \|F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1})\|\right]\ \mathfrak{f}_\alpha^{n-1} + \sin\left[\Delta t \|F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1})\|\right]\ \frac{F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1})}{\|F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1})\|}. \tag{4.3}$$

It means that $\mathfrak{f}_\alpha^{n-1}$ marches for distance $\Delta t\|F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1})\|$ along the big circle (i.e., the geodesic) determined by $\mathfrak{f}_\alpha^{n-1}$ and the flow vector $F_\alpha^{\mathrm{TV}}(\mathfrak{f}^{n-1})$. Therefore, the marching always stays on the sphere. It is easy to see that for small Δt, the two-step classical scheme is equivalent to our geodesic marching up to the leading order. Precisely, the latter always marches a distance $O(\Delta t^2)$ farther than the former (see Figure 5). The geodesic marching scheme is intrinsic for the geometry of the feature manifold, and thus it has the most general meaning for numerical restorations of nonflat features.

4.2.2. Fixed-point iteration, filtering, and the convex cone theorem. Again take the system of TV equations

$$F_\alpha^{\mathrm{TV}}(\mathfrak{f}) = \Pi_{\mathfrak{f}_\alpha}\left(\sum_{\beta\in N_\alpha} w_{\alpha\beta}\mathfrak{f}_\beta + \lambda\mathfrak{f}_\alpha^{(0)}\right) = 0, \quad \alpha\in\Omega_n,$$

for example. It can be interpreted as the following: for the optimal restoration $\mathfrak{f}$, at each node α,

$$\sum_{\beta\in N_\alpha} w_{\alpha\beta}\ \mathfrak{f}_\beta + \lambda\mathfrak{f}_\alpha^{(0)}$$

is *parallel* to $\mathfrak{f}_\alpha$. This explanation inspires the iteration scheme

$$\tilde{\mathfrak{f}}_\alpha^n = \sum_{\beta\in N_\alpha} w_{\alpha\beta}\ (\mathfrak{f}^{n-1})\mathfrak{f}_\beta^{n-1} + \lambda\mathfrak{f}_\alpha^{(0)},$$
$$\mathfrak{f}_\alpha^n = \tilde{\mathfrak{f}}_\alpha^n/\|\tilde{\mathfrak{f}}_\alpha^n\|.$$

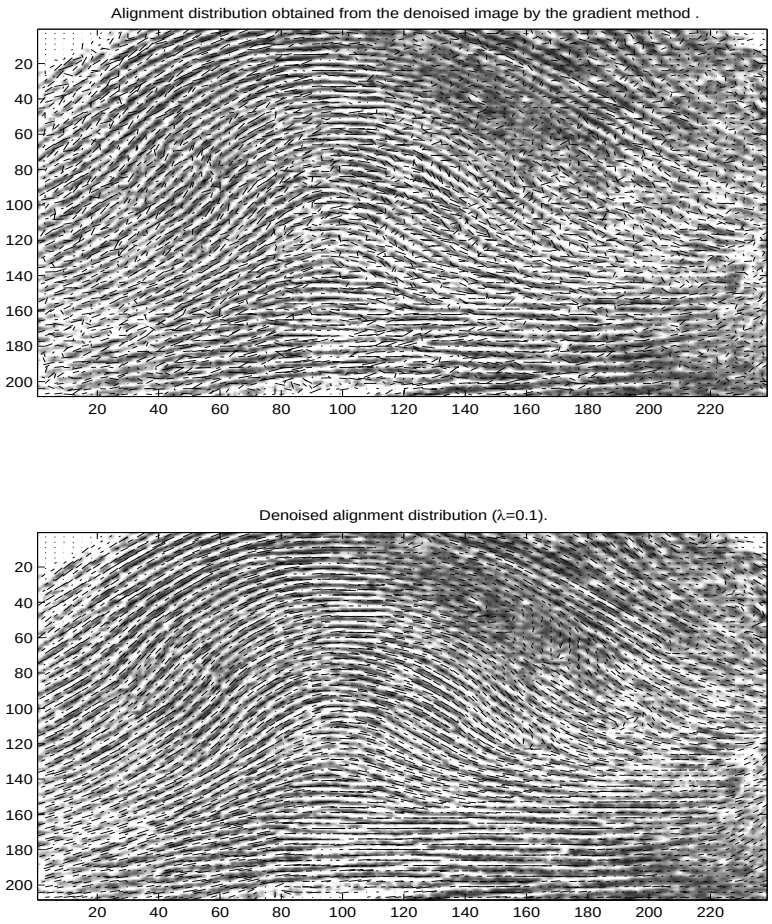

FIG. 6. *Restoring the alignment feature of fingerprints by the l^2 model equation* (3.3). *(See section* 5.1*.)*

Or, to stabilize the algorithm (and even lead to new insights), define

$$h_{\alpha\beta}(\mathfrak{f}) = \frac{w_{\alpha\beta}(\mathfrak{f})}{\sum_{\gamma\in N_\alpha} w_{\alpha\gamma}(\mathfrak{f}) + \lambda}, \qquad h_{\alpha\alpha}(\mathfrak{f}) = \frac{\lambda}{\sum_{\gamma\in N_\alpha} w_{\alpha\gamma}(\mathfrak{f}) + \lambda}.$$

We consider h as an adaptive lowpass filter. Then the above algorithm stabilizes to

$$\tilde{\mathfrak{f}}^n_\alpha = \sum_{\beta\in N_\alpha} h^{n-1}_{\alpha\beta}\mathfrak{f}^{n-1}_\beta + + h^{n-1}_{\alpha\alpha}\mathfrak{f}^{(0)}_\alpha, \tag{4.4}$$

$$\mathfrak{f}^n_\alpha = \tilde{\mathfrak{f}}^n_\alpha / \|\tilde{\mathfrak{f}}^n_\alpha\|. \tag{4.5}$$

Here $h^{n-1}_{\alpha\beta} = h_{\alpha\beta}(\mathfrak{f}^{n-1})$.

This filtering formulation immediately rewards us with one important property. A region Γ on $\mathbf{S}^2$ is said to be a *convex cone* if its spanning cone $\tilde{\Gamma}$ in $\mathbb{R}^3$,

$$\tilde{\Gamma} = \mathbb{R}^+ \times \Gamma = \{r \times \mathbf{v} : r \geq 0, \mathbf{v} \in \Gamma\},$$

is convex.

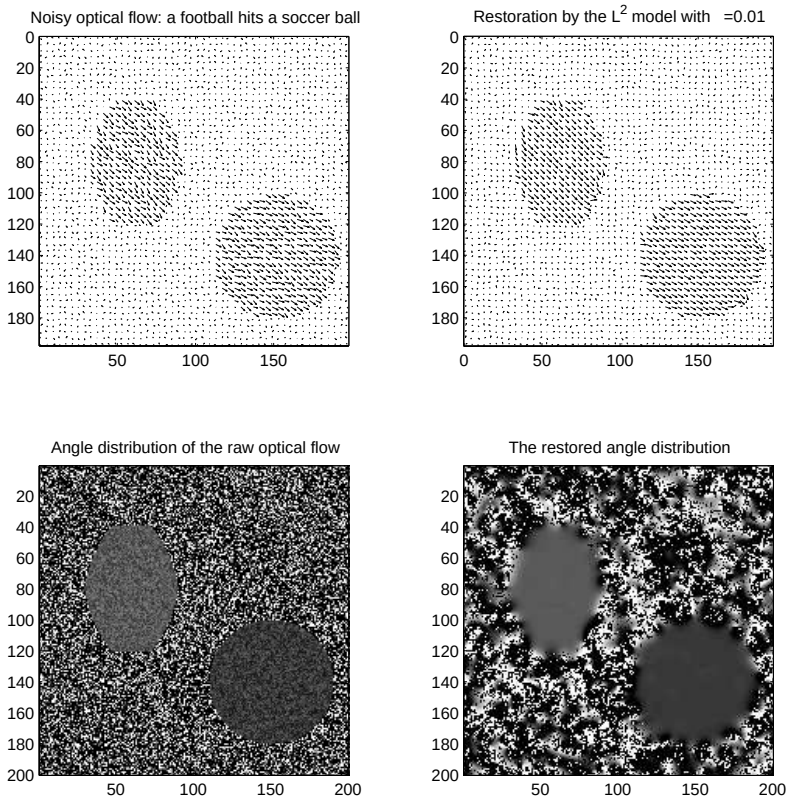

FIG. 7. *Restoration of optical flows by the* l^2 *model equation* (3.1). *(See section* 5.2.*)*

THEOREM 4.1 (convex cone theorem). *Suppose the raw data* $\mathfrak{f}_\alpha^{(0)}$*'s belong to a convex cone* Γ. *If we start the above iterations using* $\mathfrak{f}^{(0)}$ *as the initial data, then for any step* n, *all* $\mathfrak{f}_\alpha^n$*'s belong to* Γ.

The simple proof is left for our readers. Since RGB values are nonnegative, the chromaticity feature lives on the positive cone (one-eighth of the sphere), which is convex. Therefore, the above fixed-point iteration scheme keeps new feature points from wandering off the positive cone, another advantage of the discrete model.

5. Numerical examples and applications. We demonstrate the applications of our models in three important cases: fingerprint processing, optical flow restoration, and chromaticity denoising.

5.1. Restoring wave lines of a fingerprint (Figure 6). For fingerprints, the alignment pattern of wave lines is the most important feature. Therefore, instead of (3.1), we use (3.3) for restoration.

Enhancement and diffusion techniques are first applied to the original raw fingerprint image to smooth it. Then we compute the gradient flow, followed by a 90-degree rotation, to get the tangent vectors along the level lines. Removing the "heads" of all tangent vectors yields the alignment distribution for the image. The alignment pattern has been superimposed on the fingerprint image in the top subplot. Notice that it is indeed a noisy distribution regardless of our smoothing effort. In the bottom subplot, we apply the model equation (3.3) with $\lambda = 0.1$ to denoise the above distri-

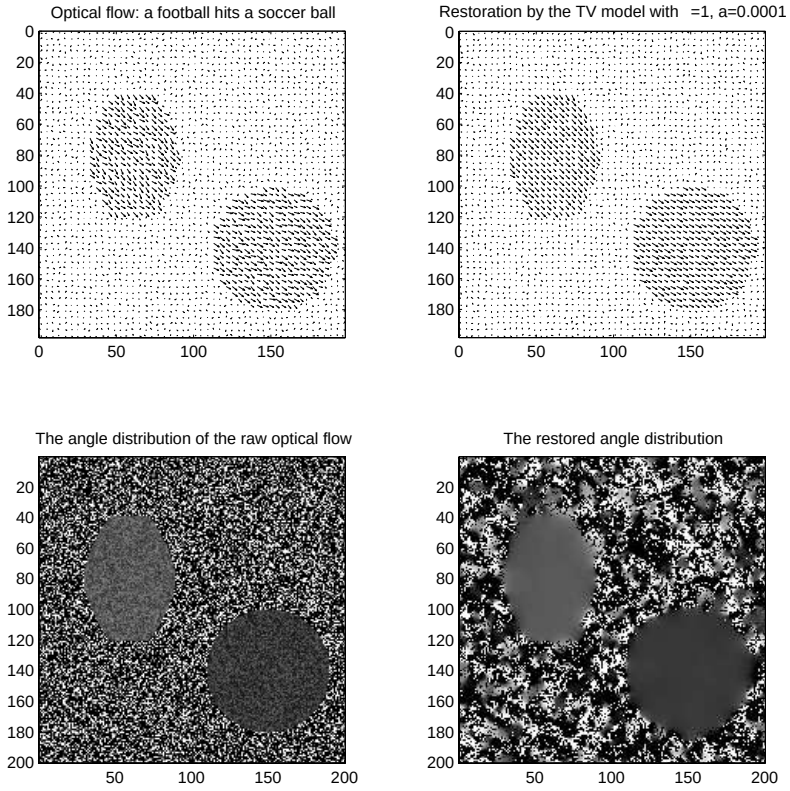

FIG. 8. *Restoration of optical flows by the TV model equation* (3.2). *(See section* 5.3.*)*

bution. The result is very successful. The alignment pattern coincides very well with the fingerprint except on a couple of local regions where the original image is blurred or damaged.

5.2. Restoration of optical flows by the l^2 model (Figure 7). In Figure 7, we show an application of the l^2 model for denoising optical flows.

The upper left subplot shows the noisy optical flow of an imagined video recording of the spinless rigid motion of a football and a soccer ball. The background is noisy.

The upper right subplot shows the restored optical flow obtained by the l^2 model. We have chosen $\lambda = 0.01$. The length of the vector (or *speed* in terms of velocity) at each pixel is not changed, and only the orientation component is denoised and restored. The two subplots at bottom plot the corresponding angle distributions for better visualization.

Observe that the motion near the boundaries is smeared due to the smoothing effect of the l^2 energy.

5.3. Restoration of optical flows by the TV model (Figure 8) . Similar to Figure 7, in Figure 8, we have tested the TV model with $\lambda = 1$. The TV model recovers the boundaries much better than the l^2 model, as is clearly observable from the angle representation.

5.4. Restoration of chromaticity (Figure 9). In Figure 9, we apply the discrete chromaticity restoration model for a clown image with Gaussian noise (the

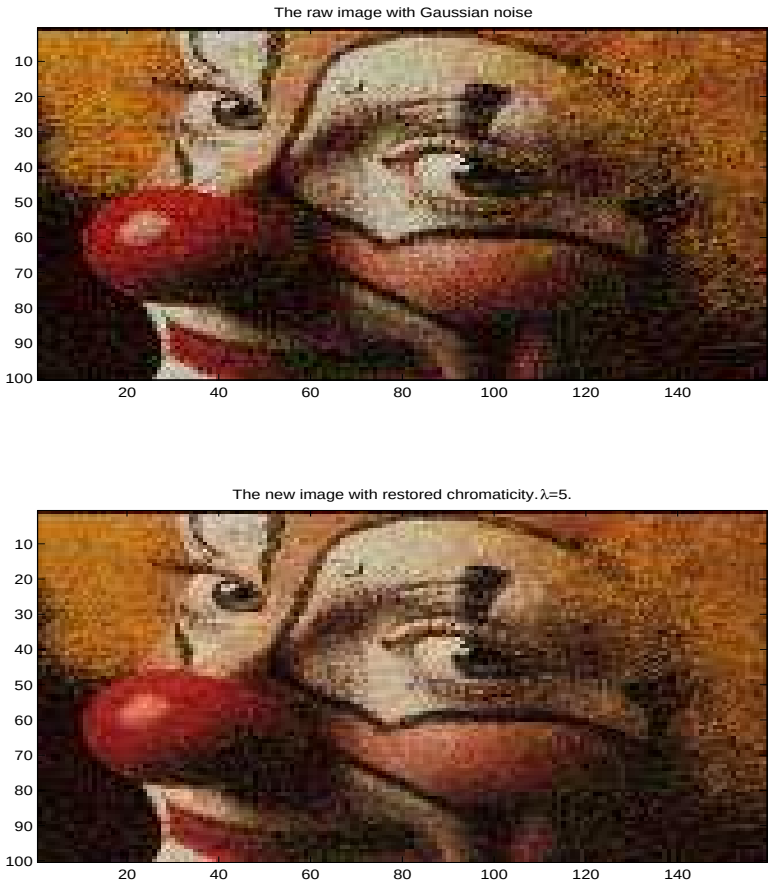

FIG. 9. *Restoration of chromaticity from Gaussian noise by the discrete TV model and spherical filtering equation* (4.4). *The brightness component remains unchanged. (See section* 5.4.*)*

top image). We first separate the vector RGB image $\mathbf{I}$ into the brightness component $B = \|\mathbf{I}\|$, and the chromaticity component $\mathfrak{f}^{(0)} = \mathbf{I}/B$. Then we apply the fixed-point TV filtering algorithm (4.4) to $\mathfrak{f}^{(0)}$ and get the optimal restoration $\mathfrak{f}$. Finally, we use the original brightness value B to assemble the new image $\mathbf{I}_{\text{new}} = \mathfrak{f} \times B$, which is shown in the bottom subplot. (One can also apply the classical flat restoration model to the scalar field B along with the chromaticity restoration. See Tang, Sapiro, and Caselles [18].) The restoration is quite successful. The original red and green dots have been almost wiped out. The eyes and dark lines resume their original black color, and the nose and lips become smoothly red.

5.5. Concluding remarks. We have developed the general mathematical models for denoising and restoration of nonflat image features. The restoration models for three types of important nonflat features are studied in detail. They are the orientation feature, the alignment feature, and the chromaticity feature. Algorithms resolving the nonlinearity of the models are constructed. Numerical evidences show that generally the TV type of models perform better than the L^2 models for nonflat features with "edges."

The current problem has close connection to harmonic maps and liquid crystals.

Interested readers can find more discussion in Perona [11] and Tang, Sapiro, and Caselles [17, 18].

Acknowledgments. It is a great pleasure to thank Professor Perona for bringing his work to our attention and for helpful discussions on the subject. We are also enormously grateful to Professors Sapiro and Osher for their numerous discussions, suggestions, and generous help. The application on chromaticity in this final version is impossible without the inspiration from Sapiro's group.

REFERENCES

[1] P. V. Blomgren and T. F. Chan, *Modular Solvers for Constrained Image Restoration*, Tech. Report CAM 97-52, Department of Mathematics, UCLA, Los Angeles, CA, 1997.
[2] A. Chambolle and P.-L. Lions, *Image recovery via total variational minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
[3] T. Chan, S. Osher, and J. Shen, *The digital TV filter and non-linear denoising*, Tech. Report CAM 99-34, Department of Mathematics, UCLA, Los Angeles, CA, 1999; IEEE Trans. Image. Process., to appear.
[4] T. Chan and L. Vese, *Variational image restoration and segmentation models and approximations*, IEEE Trans. Image Process., to appear.
[5] F. R. Chung, *Spectral Graph Theory*, CBMS Reg. Conf. Ser. Math. 92, AMS, Providence, RI, 1997.
[6] S. Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press, New York, 1962.
[7] A. Marquina and S. Osher, *Explicit Algorithms for a New Time Dependent Model Based on Level Set Motion for Nonlinear Deblurring and Noise Removal*, Tech. Report CAM 99-5, Department of Mathematics, UCLA, Los Angeles, CA, 1999.
[8] J.-M. Morel and S. Solimini, *Variational Methods in Image Segmentation*, Progr. Nonlinear Differential Equations Appl. 14, Birkhäuser, Boston, 1995.
[9] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–685.
[10] S. Osher, private communication, UCLA, Los Angeles, CA, 1999.
[11] P. Perona, *Orientation diffusion*, IEEE Trans. Image Proc., 7 (1998), pp. 457–467.
[12] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Machine Intell., 12 (1990), pp. 629–639.
[13] L. Rudin and S. Osher, *Total variation based image restoration with free local constraints*, in Proceedings of the First IEEE International Conference on Image Processing, 1 (1994), pp. 31–35.
[14] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[15] G. Sapiro, *Color Snakes*, Tech. Report HPL-95-113, Hewlett Packard Computer Peripherals Laboratory, Palo Alto, CA, 1995.
[16] G. Sapiro and D. Ringach, *Anisotropic diffusion of multivalued images with applications to color filtering*, IEEE Trans. Image Process., 5 (1996), pp. 1582–1586.
[17] B. Tang, G. Sapiro, and V. Caselles, *Color Image Enhancement via Chromaticity Diffusion*, Tech. Report, Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, 1999.
[18] B. Tang, G. Sapiro, and V. Caselles, *Direction Diffusion*, Int. J. Comput. Vision, 36 (2000), pp. 149–161.
[19] P. E. Trahanias, D. Karako, and A. N. Venetsanopoulos, *Directional processing of color images: Theory and experimental results*, IEEE Trans. Image Process., 5 (1996), pp. 868–880.
[20] P. E. Trahanias and A. N. Venetsanopoulos, *Vector directional filters—a new class of multichannel image processing filters*, IEEE Trans. Image Process., 2 (1993), pp. 528–534.
[21] C. R. Vogel and M. E. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput., 17 (1996), pp. 227–238.
[22] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner-Verlag, Stuttgart, Germany, 1998.